The evolutionary advantage of cooperation

Ole Peters[1,2,*] and Alexander Adamou[1,†]
[1]London Mathematical Laboratory, 14 Buckingham Street, London, WC2N 6DF, UK.
[2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, 87501 NM, USA.
[*]o.peters@lml.org.uk, [†]a.adamou@lml.org.uk

They give that they may live, for to withhold is to perish.
K. Gibran

June 11, 2015

Abstract

The present study asks how cooperation and consequently structure can emerge in many different evolutionary contexts. Cooperation, here, is a persistent behavioural pattern of individual entities pooling and sharing resources. Examples are: individual cells forming multicellular systems whose various parts pool and share nutrients; pack animals pooling and sharing prey; families firms, or modern nation states pooling and sharing financial resources. In these examples, each atomistic decision, at a point in time, of the better-off entity to cooperate poses a puzzle: the better-off entity will book an immediate net loss – why should it cooperate? For each example, specific explanations have been put forward. Here we point out a very general mechanism – a sufficient null model – whereby cooperation can evolve. The mechanism is based the following insight: natural growth processes tend to be multiplicative. In multiplicative growth, ergodicity is broken in such a way that fluctuations have a net-negative effect on the time-average growth rate, although they have no effect on the growth rate of the ensemble average. Pooling and sharing resources reduces fluctuations, which leaves ensemble averages unchanged but – contrary to common perception – increases the time-average growth rate for each cooperator.

At a recent Santa Fe Institute meeting it was remarked that the most striking feature of human life on earth, especially of economic life, is cooperation at all imaginable scales of organization. We would add that it is not only human life but life in general that displays cooperation. Life is full of structure – we living beings are not minimal self-reproducing chemical units, but cells, organisms, families, herds, companies, institutions, nation states and so on. This type of structure, some form of cooperation, is so widespread that we don't often ask where it comes from. Economics should be the place to look for an explanation of human social structure, but oddly the basic message from mainstream economics seems to be that optimal, rational, sensible behaviour would shun cooperation. In many ways we see cooperation in the world despite, not because of, economic theory.

Many economists are aware of this shortcoming of their discipline and are addressing it, often from psychological or neurological perspectives, as well as with the help of agent-based evolutionary simulations.

We show that cooperation and social structure arise from simple analytically solvable mathematical models for economically optimal behaviour. We contend that where economics uses models of such simplicity it ignores essential insights of the last two centuries of mathematics. Specifically, economics uses inappropriate mathematical representations of randomness. These representations have been essentially unchanged since the 17th century. As a consequence effects of fluctuations and risk (or of dynamics and time) are not properly accounted for.

In the following we introduce a minimal model of cooperation and analyze it using modern mathematical tools. This reveals an evolutionary advantage of entities that cooperate over those that do not.

Let $x_i(t)$ represent the resources of individual i at time t. For example, we could think of these resources as "calories," while being mindful that we're really exploring a mathematical model that will be more or less applicable to different real-world cases. Calories can be deployed to generate more of themselves – for instance, a tired starving wolf cannot catch prey, but a healthy strong wolf can. "Catching prey using calories" in this context means calories generating more of themselves.

A common model for such dynamics of self-reproduction is geometric Brownian motion, where $x(t)$ follows the process

$$dx = x(\mu dt + \sigma dW), \tag{1}$$

with μ the drift term, σ the amplitude of the noise, and dW a Wiener increment. The distribution of the exponential growth rates observed in an ensemble of trajectories after time T is Gaussian, and the model is an attractor for more complex models that exhibit multiplicative growth [1, 2].

The expectation value (or ensemble average) of (Eq. 1) is easily computed, $\langle dx \rangle = \langle x \rangle \mu dt$, which describes exponential growth of $\langle x \rangle$ at a rate equal to the drift term, $g_{\langle\rangle}(x) = \mu$. Naïvely, one might guess that the growth rate observed in an individual trajectory, $g(x(T)) = \frac{1}{T} \ln \left(\frac{x(T)}{x(0)} \right)$ will converge to the same value in the long time limit $\lim_{T\to\infty}$, but this is not the case. Instead, the growth rate observed in a single trajectory converges to $g_t(x) = \mu - \frac{\sigma^2}{2}$. The inequality of g_t and $g_{\langle\rangle}$, discussed in the context of ergodicity breaking in [4], is the aspect of (Eq. 1) that is of interest for the evolution of cooperation.

Having established these properties of (Eq. 1) we now introduce our model of cooperation. We start with two non-cooperating entities whose resources follow (Eq. 1) with identical drift and noise amplitude but different realizations of the noise, *i.e.* $dx_1 = x_1(\mu dt + \sigma dW_1)$ and $dx_2 = x_2(\mu dt + \sigma dW_2)$, where dW_1 and dW_2 are independent Wiener terms.

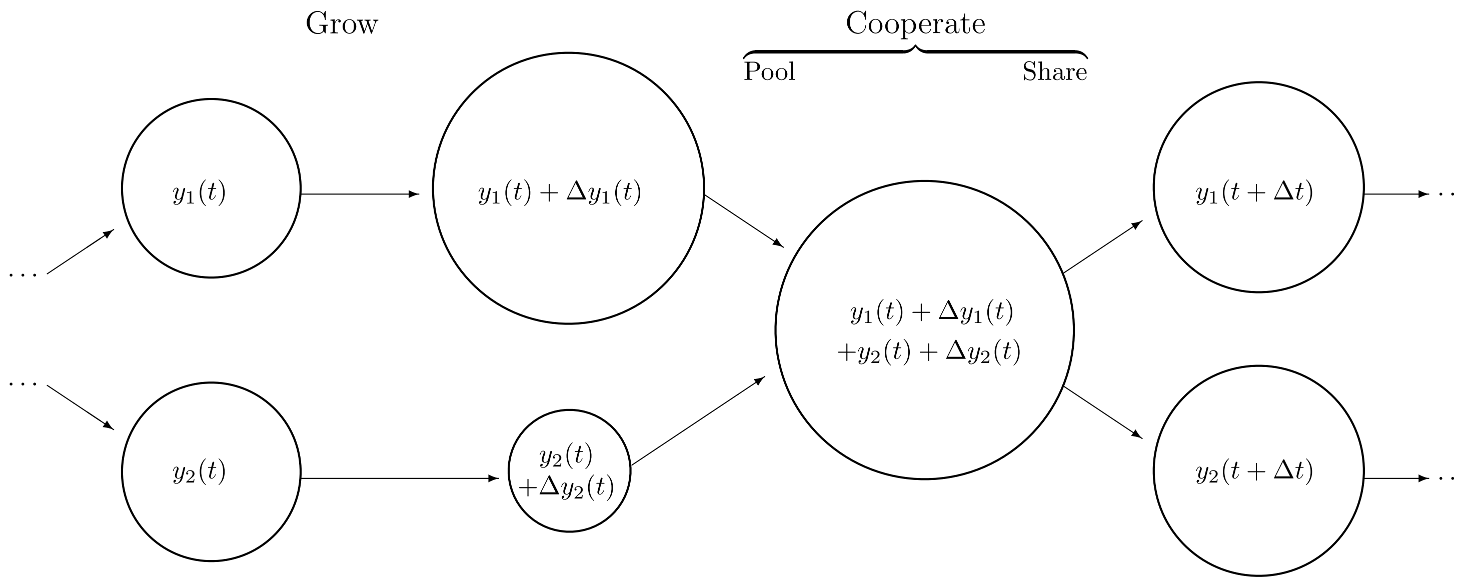

Figure 1: Cooperation dynamics. Cooperators start each time step with equal resources, then they *grow* independently according to (Eq. 5), then they *cooperate* by *pooling* resources and *sharing* them equally, then the next time step begins.

Consider a discretized version of (Eq. 1), such as would be used in a numerical simulation. The non-cooperators grow according to

$$\Delta x_i(t) = x_i(t)\left[\mu \Delta t + \sigma \sqrt{\Delta t}\, \xi_i\right], \tag{2}$$
$$x_i(t + \Delta t) = x_i(t) + \Delta x_i(t), \tag{3}$$

where ξ_i are standard normal random variates, $\xi_i \sim \mathcal{N}(0, 1)$.

Cooperation works as follows. We imagine that the two previously non-cooperating entities, with resources $x_1(t)$ and $x_2(t)$, cooperate to produce two entities, whose resources we label $y_1(t)$ and $y_2(t)$ to distinguish them from the non-cooperating case. We envisage equal sharing of resources, $y_1 = y_2$, and introduce a cooperation operator, $\oplus$, such that

$$x_1 \oplus x_2 = y_1 + y_2. \tag{4}$$

In the discrete-time picture, each time step involves a two-phase process. First there is a growth phase, analogous to (Eq. 2), in which each cooperator increases its resources by

$$\Delta y_i(t) = y_i(t)\left[\mu \Delta t + \sigma \sqrt{\Delta t}\, \xi_i\right]. \tag{5}$$

This is followed by a cooperation phase, replacing (Eq. 3), in which resources are pooled and shared equally among the cooperators:

$$y_i(t + \Delta t) = \frac{y_1(t) + \Delta y_1(t) + y_2(t) + \Delta y_2(t)}{2}. \tag{6}$$

With this prescription both cooperators and their sum experience the following dynamic:

$$(x_1 \oplus x_2)(t + \Delta t) = (x_1 \oplus x_2)(t)\left[1 + \left(\mu \Delta t + \sigma \sqrt{\Delta t}\, \frac{\xi_1 + \xi_2}{2}\right)\right]. \tag{7}$$

For ease of notation we define

$$\xi_{1\oplus 2} = \frac{\xi_1 + \xi_2}{\sqrt{2}}, \tag{8}$$

which is another standard Gaussian, $\xi_{1\oplus 2} \sim \mathcal{N}(0,1)$. Letting the time increment $\Delta t \to 0$ we recover an equation of the same form as (Eq. 1) but with a different fluctuation amplitude,

$$d(x_1 \oplus x_2) = (x_1 \oplus x_2)\left(\mu dt + \frac{\sigma}{\sqrt{2}} dW_{1\oplus 2}\right). \tag{9}$$

The expectation values of a non-cooperator, $\langle x_1(t)\rangle$, and a corresponding cooperator, $\langle y_1(t)\rangle$, are identical. From this perspective there is no incentive for cooperation. Worse still, immediately after the growth phase, the better-off entity of a cooperating pair, $y_1(t_0) > y_2(t_0)$, say, would increase its expectation value from $\frac{y_1(t_0)+y_2(t_0)}{2}\exp(\mu(t-t_0))$ to $y_1(t_0)\exp(\mu(t-t_0))$ by breaking the cooperation. An analysis based on expectation values thus finds that there is no reason reason for cooperation to arise, and that if it does arise there are good reasons for it to end, *i.e.* it will be fragile. For this reason the observation of widespread cooperation constitutes a conundrum.

The solution of the conundrum comes from considering the time average growth rate. The non-cooperating entities grow at $g_t(x_i) = \mu - \frac{\sigma^2}{2}$, whereas the cooperating unit grows at $g_t(x_1 \oplus x_2) = \mu - \frac{\sigma^2}{4}$, and we have

$$g_t(x_1 \oplus x_2) > g_t(x_i) \tag{10}$$

for any non-zero noise amplitude. A larger time-average growth rate is the textbook definition of an evolutionary advantage. The effect is illustrated in Fig. 2 by direct simulation of (Eq. 2)–(Eq. 3) and (Eq. 7).

Notice the nature of the Monte-Carlo simulation in Fig. 2. No ensemble is constructed. Only individual trajectories are simulated and run for a time that is long enough for statistically significant features to rise above the noise. This method teases out of the dynamics what happens over time. The significance of any observed structure – its epistemological meaning – is immediately clear: this is what happens over time for an individual system (a cell, a person's wealth, *etc.*). Simulating an ensemble and averaging over members to remove noise does not tell the same story. The resulting features may not emerge over time. They are what happens on average in an ensemble, but this may have no bearing on what happens to the individual. For instance the pink dashed line in Fig. 2 is the ensemble average of $x_1(t)$, $x_1(t)$, and $(x_1 \oplus x_2)(t)/2$, and it has nothing to do with what happens in the individual trajectories over time.

In our model the advantage of cooperation, and hence the emergence of structure is purely a non-linear effect of fluctuations – cooperation reduces the magnitude of fluctuations, and over time (though not in expectation) this implies faster growth.

A simple extension of the model is the inclusion of more cooperators. For n cooperators, the spurious drift term is $-\frac{\sigma^2}{2n}$, so that the time-average growth approaches expectation-value growth for large n. When judged on expectation values, the apparent futility of cooperation is unsurprising because expectation values are the result for infinitely many cooperators, and adding further cooperators cannot improve on this.

The approach to this upper bound as the number of cooperators increases favours the formation of social structure. We may generalise to different drift terms, μ_i, and noise amplitudes, σ_i, for different individual entities. Again it is straightforward to see

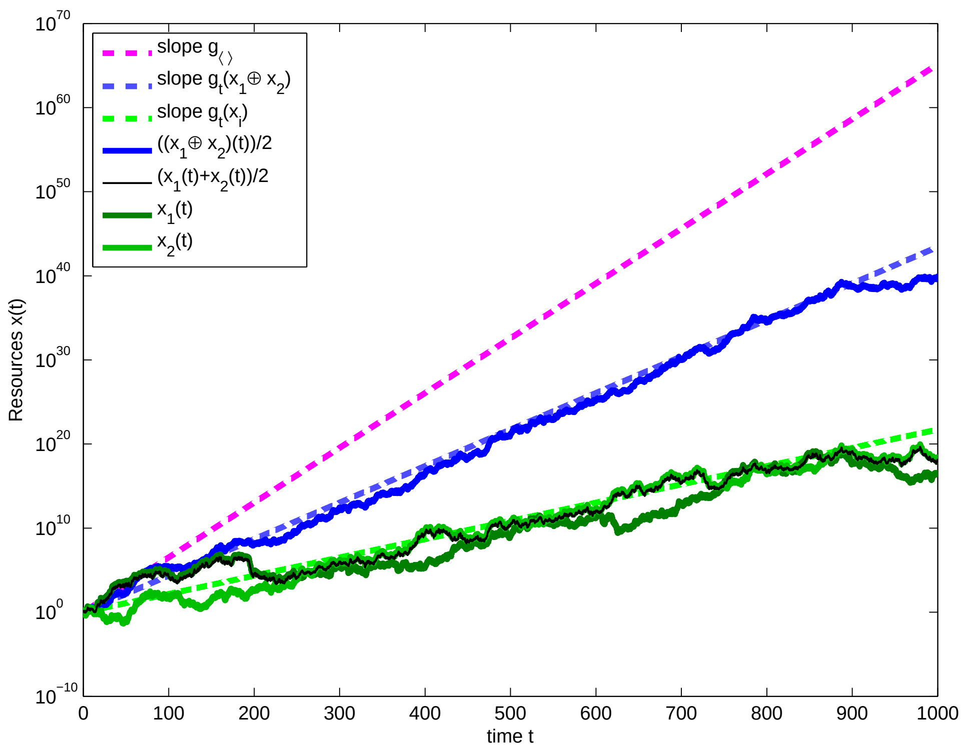

Figure 2: Typical trajectories for two non-cooperating (green) entities and for the corresponding cooperating unit (blue). Over time, the noise reduction for the cooperator leads to faster growth. Even without effects of specialisation or the emergence of new function, cooperation pays in the long run. The black thin line shows the average of the non-cooperating entities. While in the logarithmic vertical scale the average traces the more successful trajectory, it is far inferior to the cooperating unit. In a very literal mathematical sense the whole, $(x_1 \oplus x_2)(t)$, is more than the sum of its parts, $x_1(t) + x_2(t)$. The algebra of cooperation is not merely that of summation.

whether cooperation is beneficial in the long run for any given entity. An entity whose drift term, μ_1, is larger than a potential cooperator's drift term, μ_2, still benefits from cooperation if

$$\mu_1 - \frac{\sigma_1^2}{2} < \frac{\mu_1 + \mu_2}{2} - \frac{\sigma_1^2 + \sigma_2^2}{8}. \tag{11}$$

Another generalisation is partial cooperation – entities can be assumed to share only a proportion of their resources, resembling taxation and redistribution. We discuss this in a separate manuscript.

In reality many other effects contribute to the formation of structure, such as the emergence of new function and the possibility of specialisation. An example of new function is the ability of relatively small multicellular organisms to swim or funnel nutrients towards themselves [5, 6], which is necessary to satisfy nutrient demand for relatively smaller reaction surfaces. Examples of specialisation are found in larger organisms, with different cell types performing specialised tasks. They are also found in many forms of human cooperation, with firms or individuals becoming proficient in specific tasks.

The impact of risk reduction on time-average growth suggests that risk management has a rarely recognised significance. Fluctuation reduction (*i.e.* good risk management) does not merely reduce the likelihood of disaster or the size of up and down swings but also it improves the time-average performance of the structure whose risks are being managed. In a financial context the value of a portfolio whose risks are well managed will not just display smaller fluctuations, but will grow faster in time. Similarly, a well-diversified economy will grow faster in the long run than a poorly diversified economy.

How significant or insignificant the effects of fluctuation-reduction are depends on the specific setup, but it is tantalising to see that the simple universal setting of multiplicative growth generates structure, *i.e.* large units of cooperation.

We have explored here momentous consequences of the difference between $g_{\langle\rangle}$ and g_t. A property as mathematically simple and important to our understanding of structure, especially observed in economic systems, should be discussed on page 1 of any textbook on economics, ecology, evolutionary biology. For historical reasons this is not the case. The first formal treatments of randomness were motivated by gambling problems – in 1654 Fermat and Pascal asked how to split the pot in a fair manner for an unfinished game and invented the expectation value [7]. From this grew utility theory with Daniel Bernoulli's treatment of the St. Petersburg gamble in 1738 [8], which forms the basis of all of modern formal economics, including finance and macroeconomics, according to Cochrane [9]. An unbroken conceptual line can be drawn from 17th and 18th century probability theory to modern economics.

The insight that time averages may not be identical to expectation values was only reached in the development of statistical mechanics in 19th century physics, where the physical nature of expectation values was elucidated as the mean over an ensemble of non-interacting systems. The development of ergodic theory in the 20th and 21st centuries provided the concepts that reveal the physically relevant nature of stochastic processes like (Eq. 1).

Judging from our experience, these developments were not adopted by mainstream economics. Economics was the first scientific discipline to embrace randomness as an important descriptor of the physical world. Unfortunately during this trailblazing immature mathematical concepts became engrained in the formalism.

Today, concepts that ignore basic non-linear dynamical effects such as those we have described here form the curriculum of economics and are being spread into economic policy making. Communicating the inappropriateness of these concepts is

difficult – the vocabulary required to describe the problem is not that of present-day economics.

It may be a generation-defining challenge to develop an economic formalism that takes as its basis the mathematical properties we have discussed here. In the process we will learn whether our natural tendency to cooperate, our gut feeling and moral sentiment, is in harmony with a careful formal analysis of the issues involved.

References

[1] Aitchison, J. & Brown, J. A. C. *The lognormal distribution* (Cambridge University Press, 1957).

[2] Redner, S. Random multiplicative processes: An elementary tutorial. *Am. J. Phys.* **58**, 267–273 (1990).

[3] Kloeden, P. E. & Platen, E. *Numerical solution of stochastic differential equations*, vol. 23 (Springer Science & Business Media, 1992).

[4] Peters, O. & Klein, W. Ergodicity breaking in geometric Brownian motion. *Phys. Rev. Lett.* **110**, 100603 (2013).

[5] Short, M. B. *et al.* Flows driven by flagella of multicellular organisms enhance long-range molecular transport. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8315–8319 (2006).

[6] Roper, M., Dayel, M. J., Pepper, R. E. & Koehl, M. A. R. Cooperatively generated stresslet flows supply fresh fluid to multicellular choanoflagellate collonies. *Phys. Rev. Lett.* **110**, 228104 (2013).

[7] Devlin, K. *The unfinished game* (Basic Books, 2008).

[8] Bernoulli, D. Specimen Theoriae Novae de Mensura Sortis. Translation "Exposition of a new theory on the measurement of risk" by L. Sommer (1954). *Econometrica* **22**, 23–36 (1738). URL `http://www.jstor.org/stable/1909829`.

[9] Cochrane, J. H. *Asset Pricing* (Princeton University Press, 2001).